

# Statistical comparison of ensemble implementations of Grover's search algorithm to classical sequential searches.

Tomasz M. Kott

Department of Physics, University of Maryland, College Park, MD 20742 USA\*

David Collins

Department of Physical and Environmental Sciences, Mesa State College,  
1100 North Avenue, Grand Junction, CO 81501, USA†

We compare pseudopure state ensemble implementations, quantified by their initial polarization and ensemble size, of Grover's search algorithm to probabilistic classical sequential search algorithms in terms of their success and failure probabilities. We propose a criterion for quantifying the resources used by the ensemble implementation via the aggregate number of oracle invocations across the entire ensemble and use this as a basis for comparison with classical search algorithms. We determine bounds for a critical polarization such that the ensemble algorithm succeeds with a greater probability than the probabilistic classical sequential search. Our results indicate that the critical polarization scales as  $N^{-1/4}$  where  $N$  is the database size and that for typical room temperature solution state NMR, the polarization is such that the ensemble implementation of Grover's algorithm would be advantageous for  $N \gtrsim 10^{22}$ .

PACS numbers: 03.67.Lx

## I. INTRODUCTION

Conventional implementations of quantum algorithms entail application of unitary transformations and projective measurements to a single, multi-qubit quantum system which is initially prepared in a known pure state [1, 2, 3, 4]. In the alternative ensemble paradigm, of which solution state nuclear magnetic resonance (NMR) is the most prominent example, the algorithm is implemented on an ensemble of multiple identical, noninteracting quantum systems [5, 6, 7, 8, 9, 10, 11, 12, 13, 14].

The principal differences between the two paradigms manifest themselves in the preparation, measurement and post-measurement stages and arise from: (i) the impossibility of isolating any given ensemble members and applying preparation, evolution and measurement operations to these alone while ignoring the remaining ensemble members and (ii) the fact that available initial states are mixed. The principal similarity between the paradigms is that it is possible to apply identical unitary evolution operators to each ensemble member at any stage of the algorithm. A variety of schemes have been developed to deal with the initialization, measurement and algorithm output issues in ensemble implementations of quantum algorithms [6, 7, 15, 16, 17, 18]. Collectively these issues are resolved via the use of a *pseudopure initial state*, for whose preparation there are several known schemes [5, 6, 15, 16, 17], described by the density oper-

ator

$$\hat{\rho}_i = \frac{(1-\varepsilon)}{2^n} \hat{I}^{\otimes n} + \varepsilon |\psi_i\rangle\langle\psi_i| \quad (1)$$

where  $n$  is the number of qubits,  $|\psi_i\rangle$  is the initial pure state required by the conventional (single quantum system) implementation of the algorithm and  $0 \leq \varepsilon \leq 1$  is called the *polarization*. Under the unitary transformation for the algorithm,  $\hat{U}_{\text{alg}}$ , the density operator transforms to

$$\begin{aligned} \hat{\rho}_{\text{final}} &= \frac{(1-\varepsilon)}{2^n} \hat{I}^{\otimes n} + \varepsilon \hat{U}_{\text{alg}} |\psi_i\rangle\langle\psi_i| \hat{U}_{\text{alg}}^\dagger \\ &= \frac{(1-\varepsilon)}{2^n} \hat{I}^{\otimes n} + \varepsilon |\psi_{\text{final}}\rangle\langle\psi_{\text{final}}| \end{aligned} \quad (2)$$

where

$$|\psi_{\text{final}}\rangle := \hat{U}_{\text{alg}} |\psi_i\rangle \quad (3)$$

is the final, pre-measurement state of the quantum system in the conventional implementation of the algorithm. This gives the appearance of a conventional realization of the algorithm on a subset of the ensemble which is initially described by the density operator  $|\psi_i\rangle\langle\psi_i|$ , which corresponds to a pure state. In most proposals for and realizations of ensemble implementations the algorithm output is extracted from the expectation values of single qubit traceless observables [6, 7, 17, 18], as we describe in Sec. III. Such expectation values are deterministic. However, in real implementations the number of ensemble members is finite and expectation values can only be approximated by sample averages. Thus, regardless of the nature of the original quantum algorithm, ensemble implementations are probabilistic (except possibly in the special case where  $\varepsilon = 1$  which reduces to a conventional implementation). Although this basic feature has

\*Electronic address: tkott@umd.edu

†Electronic address: dacollin@mesastate.edu;  
URL: <http://home.mesastate.edu/~dacollin/index.html>;  
Author to whom correspondence should be addressed.

been noted [7, 19], there has been limited discussion of its implications with regard to reasonable comparisons of ensemble implementations of quantum algorithms to probabilistic classical competitors. Such issues have been addressed in the context of various ensemble implementations of single bit output algorithms such as the Deutsch-Jozsa algorithm [20, 21], resulting in bounds of the polarization beneath which a classical probabilistic algorithm will succeed with greater probability than the ensemble implementation of the quantum algorithm.

In this article we compare the performance of an ensemble implementation of the Grover search algorithm to probabilistic classical sequential searches. Our aim is to determine the polarization as a function of both ensemble and database size such that the ensemble implementation of the quantum algorithm succeeds with greater probability than its classical competitor. The primary distinction between this and previous work on the Deutsch-Jozsa algorithm is that the output of the Grover algorithm is determined from measurements on many qubits and that these measurement outcomes are correlated.

The article is organized as follows. We describe the conventional Grover search algorithm in Sec. II. In Sec. III we describe an ensemble implementation of the Grover search algorithm and explicitly provide a protocol for extracting an algorithm output. The bulk of our work is found in Sec. IV, which covers the probability with which ensemble implementations succeed and uses these as a basis for comparisons with classical probabilistic searches. Finally the appendices contain elaborations of the mathematics behind our results.

## II. GROVER'S SEARCH ALGORITHM: CONVENTIONAL IMPLEMENTATION

The simplest unstructured database search problem involves a database that conceals a single marked item in one of  $N$  possible locations, conveniently represented by the integers  $\{0, 1, \dots, N-1\}$ . The task is to find the location of the marked item,  $s$ , with the help of an oracle,

$$f(x) := \begin{cases} 0 & \text{if } x \neq s \\ 1 & \text{if } x = s. \end{cases} \quad (4)$$

A classical sequential search proceeds by evaluating the oracle at distinct database locations and on average requires  $N/2$  such oracle queries to determine the marked item's location.

The Grover search algorithm [1, 22] requires  $n = \lceil \log_2 N \rceil$  qubits whose computational basis states  $\{|x\rangle, x = 0, \dots, 2^n - 1\}$  can be regarded as representing database locations. The oracle is implemented via a unitary transformation, defined on the computational basis states as

$$\hat{U}_f |x\rangle := (-1)^{f(x)} |x\rangle \quad (5)$$

and extended linearly to all superpositions. The oracle is easily extended to the database locations  $\{N, N +$

$1, \dots, 2^n - 1\}$  by setting  $f(x) = 0$  at these points. For this reason we shall restrict discussion of the quantum algorithm to cases where  $N = 2^n$  for some integer  $n$ . The algorithm unfolds by initializing the qubits to the state

$$|\psi_i\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{N-1} |x\rangle \quad (6)$$

followed by repeated applications of the *Grover iterate*,  $\hat{G} := \hat{D}\hat{U}_f$ , where the unitary “inversion about the average” operation is

$$\hat{D} \left( \sum_{x=0}^{N-1} c_x |x\rangle \right) := \sum_{x=0}^{N-1} (-c_x + 2\langle c \rangle) |x\rangle \quad (7)$$

with  $\langle c \rangle = \sum_{x=0}^{N-1} c_x / N$ . Each application of  $\hat{G}$  includes one oracle query. For a database containing only one marked item, to which case we restrict our consideration, a standard analysis [1, 23] demonstrates that after  $q$  applications of  $\hat{G}$  the state of the system is

$$|\psi\rangle = \alpha_q |s\rangle + \beta_q \frac{1}{\sqrt{N-1}} \sum_{x \neq s} |x\rangle \quad (8)$$

where

$$\alpha_q := \sin \left( \frac{2q+1}{2} \theta \right), \quad (9)$$

$$\beta_q := \cos \left( \frac{2q+1}{2} \theta \right) \quad (10)$$

and  $0 \leq \theta \leq \pi$  depends on  $N$  via

$$\theta = \arccos \left( 1 - \frac{2}{N} \right). \quad (11)$$

The probability that the final computational basis measurement will yield  $s$  is

$$\Pr(\text{correct}) = |\langle s | \psi \rangle|^2 = \sin^2 \left( \frac{2q+1}{2} \theta \right) \quad (12)$$

and the measurement returns the marked item's location with certainty whenever  $(2q+1)\theta/2$  is an odd multiple of  $\pi/2$ . The lowest value of  $q$  which ensures this is  $\pi/(2\theta) - 1/2$  but for an arbitrary database size,  $\theta$  may not be such that this is integral. The standard protocol for the Grover algorithm requires that the number of applications of  $\hat{G}$  is

$$q_{\text{std}} = \text{nint} \left( \frac{\pi}{2\theta} - \frac{1}{2} \right) \quad (13)$$

where  $\text{nint}(x)$  is the nearest integer to  $x$ . The probability with which a computational basis measurement will yield  $s$  correctly can be ascertained by noting that  $\pi/(2\theta) - 1 \leq q_{\text{std}} \leq \pi/(2\theta)$ . At either extreme, the probability that the computational basis measurement yields

$s$  is  $\cos^2(\theta/2) = 1 - 1/N$ ; thus the probability of successfully locating the marked item after  $q_{\text{std}}$  oracle invocations is at least  $1 - 1/N$ . It can be shown that, throughout this range, the probability that the final computational basis measurement yields  $s$  is bounded from below by the probability at the extremes of the range. Typically, the search problem is considered for  $N \gg 1$ , in which case the Grover algorithm yields the marked item's location with near certainty. The number of oracle invocations required for the standard algorithm can be approximated from Eq. (11)

$$\theta \approx \frac{2}{\sqrt{N}} \left[ 1 + O\left(\frac{1}{N}\right) \right] \approx \frac{2}{\sqrt{N}} \quad (14)$$

and thus

$$q_{\text{std}} = \text{nint} \left( \frac{\pi}{4} \sqrt{N} - \frac{1}{2} \right) \approx \text{nint} \left( \frac{\pi}{4} \sqrt{N} \right). \quad (15)$$

### III. GROVER SEARCH ALGORITHM ON A ENSEMBLE OF QUANTUM SYSTEMS

We consider an ensemble implementation of Grover's algorithm starting with a pseudopure state of the type given by Eq. (1). Eqs. (2) and (8) imply that the density operator after  $q$  applications of  $\hat{G}$  and prior to measurement is

$$\hat{\rho}_{\text{final}} = \frac{(1 - \varepsilon)}{2^n} \hat{I}^{\otimes n} + \varepsilon |\alpha_q|^2 |s\rangle\langle s| + \varepsilon \frac{\alpha_q \beta_q^*}{\sqrt{N-1}} \sum_{x \neq s} |s\rangle\langle x| + \varepsilon \frac{\alpha_q^* \beta_q}{\sqrt{N-1}} \sum_{x \neq s} |x\rangle\langle s| + \varepsilon \frac{|\beta_q|^2}{N-1} \sum_{x \neq s} \sum_{y \neq s} |x\rangle\langle y|. \quad (16)$$

The standard protocol for determining the algorithm outcome, and one which eliminates the term proportional to  $\hat{I}$ , is based on the *expectation values of traceless single qubit observables* [6, 19]. For the  $j^{\text{th}}$  qubit, single qubit expectation values can be computed from  $\hat{\rho}_{\text{final red}}^{(j)}$ , the final reduced density operator for the qubit. For Grover's algorithm, Eq. (16) implies that, after  $q$  applications of  $\hat{G}$ ,

$$\begin{aligned} \hat{\rho}_{\text{final red}}^{(j)} = & \frac{1}{2} \left[ 1 - \varepsilon \frac{|\alpha_q|^2 N - 1}{N - 1} \right] \hat{I} + \varepsilon \frac{|\alpha_q|^2 N - 1}{N - 1} |s_j\rangle\langle s_j| \\ & + \varepsilon \left( \frac{N}{2} - 1 \right) \frac{1 - |\alpha_q|^2}{N - 1} \left[ |s_j\rangle\langle s_j \oplus 1| + |s_j \oplus 1\rangle\langle s_j| \right] \\ & + \varepsilon \frac{\alpha_q^* \beta_q}{\sqrt{N-1}} |s_j \oplus 1\rangle\langle s_j| + \varepsilon \frac{\alpha_q \beta_q^*}{\sqrt{N-1}} |s_j\rangle\langle s_j \oplus 1| \end{aligned} \quad (17)$$

where  $s_j$  is the  $j^{\text{th}}$  bit of  $s$ . The expectation value,

$$\begin{aligned} \langle \sigma_z \rangle^{(j)} &= \text{Tr} \left( \hat{\sigma}_z \hat{\rho}_{\text{final red}}^{(j)} \right) \\ &= (-1)^{s_j} \varepsilon \frac{|\alpha_q|^2 N - 1}{N - 1} \end{aligned} \quad (18)$$

depends on  $s_j$ . Thus  $s_j$  can be determined whenever the *expectation value of  $\hat{\sigma}_z$  can be measured for the  $j^{\text{th}}$  qubit* and the measurement resolution allows for distinction between  $+\varepsilon(|\alpha_q|^2 N - 1)/(N - 1)$  and  $-\varepsilon(|\alpha_q|^2 N - 1)/(N - 1)$ . This opens the possibility for a truncated version of Grover's algorithm, which uses fewer oracle invocations (i.e.  $q < q_{\text{std}}$ ) than required by the standard version of Grover's algorithm, subject to the proviso that  $|\alpha_q|$  is large enough to distinguish between the two possible expectation values [18]. If these requirements are satisfied, then an ensemble implementation based on expectation

values will be *deterministic* in the sense that a given value of  $s$  will always yield the same outcomes and the value of  $s$  can be determined with certainty.

For an ensemble containing a finite number of members,  $M$ , precise expectation value measurements are idealizations. Their purpose is to suggest protocols that produce algorithm outputs based on appropriate sample averages of projective measurement outcomes for individual ensemble members. The key notion is that, for a sufficiently large  $M$ , sample averages of projective measurement outcomes give close approximations to expectation values. The typical protocol has already been described for single bit output algorithms [20] and is readily extended to multiple bit output algorithms. First, a computational basis measurement is performed across all qubits for each ensemble member. Consider the  $j^{\text{th}}$  qubit of ensemble member  $k$  (where  $1 \leq k \leq M$ ). The projectors

for the two possible computational basis measurement outcomes are  $\hat{P}_{s_j} := |s_j\rangle\langle s_j|$  and  $\hat{P}_{\bar{s}_j} := |s_j \oplus 1\rangle\langle s_j \oplus 1|$  and the corresponding measurement outcome are scaled to  $z_k^{(j)} = +(-1)^{s_j}$  and  $z_k^{(j)} = -(-1)^{s_j}$  respectively; this translates between the computational basis measurement outcomes and the eigenvalues of  $\sigma_z$ . Second, the sample average for the  $j$ th qubit is computed

$$\bar{z}^{(j)} := \frac{1}{M} \sum_{k=1}^M z_k^{(j)} \quad (19)$$

and this approximates  $\langle \sigma_z \rangle^{(j)}$ . Finally, if  $\bar{z}^{(j)} > 0$  (or conversely  $\bar{z}^{(j)} < 0$ ) then this suggests that  $\langle \sigma_z \rangle^{(j)} > 0$ , (or  $\langle \sigma_z \rangle^{(j)} < 0$ ) which, according to Eq. (18) implies  $s_j = 0$  (or  $s_j = 1$ ). Thus a *candidate*,  $c := c_n \dots c_1$ , for the marked item's location is assigned via:

$$\begin{aligned} \bar{z}^{(j)} > 0 &\Rightarrow c_j = 0 \\ \bar{z}^{(j)} = 0 &\Rightarrow \begin{cases} c_j = 0 \text{ with probability } 1/2, \\ c_j = 1 \text{ with probability } 1/2 \text{ and} \end{cases} \\ \bar{z}^{(j)} < 0 &\Rightarrow c_j = 1. \end{aligned} \quad (20)$$

The protocol of Eqs. (19) and (20) is equivalent to taking a bitwise majority vote of the computational basis measurement outcomes over the entire ensemble. If, for a given bit, more ensemble members return +1 than -1 after computational basis measurements, then the intuition is that  $s_j = 0$  since the outcome  $(-1)^{s_j}$  is more likely than  $-(-1)^{s_j}$ . The string  $c := c_n \dots c_1$  constructed in this fashion constitutes the ensemble algorithm's output.

The probabilities with which single qubit outcomes occur are readily determined from the reduced density operator of Eq. (17). Specifically the probability that an ensemble member yields the "correct" value is

$$\begin{aligned} \Pr(z_k^{(j)} = +(-1)^{s_j}) &= \text{Tr}(\hat{\rho}_{\text{final red}}^{(j)} \hat{P}_{s_j}) \\ &= \frac{1 + \varepsilon_{\text{eff}}}{2} \end{aligned} \quad (21)$$

where the *effective polarization* is

$$\varepsilon_{\text{eff}} := \varepsilon \frac{|\alpha_q|^2 N - 1}{N - 1}. \quad (22)$$

Similarly the probability that an ensemble member yields the "incorrect" value is

$$\Pr(z_k^{(j)} = -(-1)^{s_j}) = \frac{1 - \varepsilon_{\text{eff}}}{2}. \quad (23)$$

Unless both  $\varepsilon = 1$  and  $|\alpha_q|^2 = 1$ , the protocol of Eqs. (19) and (20) can give  $c_j \neq s_j$ , thus causing the algorithm to fail to locate the marked item correctly. In this sense the algorithm is *probabilistic*. Even for perfect polarization, this is generally unavoidable since, for the standard number of invocations of the Grover iterate, i.e.

$q = q_{\text{std}}$ ,  $|\alpha_q|^2 \neq 1$ . However, for the standard number of iterations,

$$1 - \frac{1}{N} \leq |\alpha_q|^2 \leq 1 \quad (24)$$

giving

$$\varepsilon \left(1 - \frac{1}{N-1}\right) \leq \varepsilon_{\text{eff}} \leq \varepsilon \quad (25)$$

and as  $N \rightarrow \infty$ ,  $\varepsilon_{\text{eff}} \approx \varepsilon$ . Thus for the *typical scenario*, for which  $N \gg 1$  and  $q = q_{\text{std}}$  in ensemble implementations the probabilistic aspects will be determined predominantly by the polarization. We shall henceforth only consider the typical scenario, taking  $\varepsilon_{\text{eff}} = \varepsilon$  and  $q = \pi\sqrt{N}/4$ .

The protocol of Eq. (20) is one in which it is assumed that it is possible to distinguish a discrepancy as small as 1 between the numbers of projective measurement outcomes returning +1 and -1 respectively. This is the *best resolution case* and one can generalize the situation to consider cases where the resolution is worse as has been done for single-bit output algorithms [20]. In this article we shall restrict the discussion to the best resolution case.

#### IV. STATISTICAL PERFORMANCE OF THE ENSEMBLE QUANTUM SEARCH ALGORITHM

The probabilistic nature of quantum algorithms on ensembles motivates performance comparisons, based on the probability with which each algorithm fails or succeeds to correctly identify the marked item's location, with probabilistic classical competitors. Of the possible scenarios, we shall consider the probability of correctly identifying *every bit*  $s_1, \dots, s_n$  of the marked item's location. Eqs (21) and (22) suggest that this will depend on the database size  $N$  and the polarization  $\varepsilon$  (for the typical scenario  $q$  is determined by  $N$ ). Also, as with any sample average, it is to be expected that the failure or success probability will depend on the ensemble size  $M$ . Denote the probability with which the ensemble quantum algorithm correctly identifies *every bit of*  $s$  by  $p_{\text{success all}}^q(\varepsilon, M, N)$  and the probability with which it fails (i.e. at least one bit of  $s$  is incorrectly identified) by  $p_{\text{fail all}}^q(\varepsilon, M, N) = 1 - p_{\text{success all}}^q(\varepsilon, M, N)$ . Methods for calculating and approximating these have been provided for ensemble quantum algorithms with single bit outputs [20]. The key difference here is that Grover's algorithm's output involves multiple qubits and that measurement outcomes for *distinct qubits in an individual ensemble member are correlated*. Thus, although we can still use the methods for single bit outputs [20] to determine the probability with which any single bit value will be correctly identified, the fact that the measurement outcomes for different qubits are not statistically

independent means that success probabilities for multiple output bit cases will not necessarily be simple combinations of those for the single output bit case.

To illustrate this point, consider the case where  $n = 2$  (i.e.  $N = 4$ ) and the marked item is located at  $s = 11$  (in binary notation). Eq. (11) implies that  $\theta = \pi/3$  giving  $g_{\text{std}} = 1$ . Then Eq. (9) implies  $|\psi_f\rangle = |s\rangle = |11\rangle$ . For an ensemble realization, Eq. (22) gives  $\varepsilon_{\text{eff}} = \varepsilon$ . This yields the joint probability distribution for various measurement outcomes on both qubits of a single ensemble member of Table I.

The marginal probability distributions for single qubit outcomes are identical; for the  $k^{\text{th}}$  ensemble member and the  $j^{\text{th}}$  bit,  $\Pr(z_k^{(j)} = 0) = (1 - \varepsilon)/2$  and  $\Pr(z_k^{(j)} = 1) = (1 + \varepsilon)/2$ . Thus the probability distributions for the individual qubit outcomes are *not* independent except in the trivial cases of  $\varepsilon = 0$  or  $\varepsilon = 1$ . Let  $k_1$  denote the number of times that the outcome for qubit 1 is 1 and  $k_2$  the number of times that it is 1 for qubit 2. When the number of ensemble members is odd the ensemble algo-

rithm succeeds provided that  $k_1 > M/2$  and  $k_2 > M/2$ . It is shown in appendix A that the probability with which single qubit measurements yield  $k_1$  and  $k_2$  is

Bit 2 Outcome	Bit 1 Outcome	Probability
0	0	$\frac{1 - \varepsilon}{4}$
0	1	$\frac{1 - \varepsilon}{4}$
1	0	$\frac{1 - \varepsilon}{4}$
1	1	$\frac{1 + 3\varepsilon}{4}$

TABLE I: Joint probability distribution for  $N = 4$  and  $s = 11$  (binary).

$$\Pr(k_2, k_1) = \frac{1}{4^M} \sum_{l=\max(0, k_1+k_2-M)}^{\min(k_1, k_2)} (M - k_1 - k_2 + l, k_1 - l, k_2 - l, l)! (1 - \varepsilon)^{M-l} (1 + 3\varepsilon)^l \quad (26)$$

where  $(n_1, n_2, n_3, n_4)! = (n_1 + n_2 + n_3 + n_4)! / (n_1! n_2! n_3! n_4!)$ . Then

$$p_{\text{success all}}^q(\varepsilon, M, N) = \sum_{k_1=\frac{M+1}{2}}^M \sum_{k_2=\frac{M+1}{2}}^M \Pr(k_2, k_1). \quad (27)$$

For algorithms with single bit outputs, the expression analogous to Eq. (27) is a cumulative binomial distribution, which can be expressed as an incomplete beta function, greatly simplifying various calculations and inferences [20]. We are unaware of comparable techniques for re-expressing Eq. (27) in cases involving more than one output bit.

Nevertheless certain general conclusions are possible. If  $\varepsilon = 0$ , the ensemble is in the maximally mixed state throughout the algorithm. Thus for a given qubit, a computational basis measurement yields either outcome with equal probability. The procedure for deciding the algorithm output amounts to unbiased guessing on each bit and these outcomes are no longer correlated. It follows that

$$p_{\text{success all}}^q(0, M, N) = \frac{1}{2^n} = \frac{1}{N}. \quad (28)$$

At the other extreme  $\varepsilon = 1$  and  $\rho_{\text{final}} = |\psi_{\text{final}}\rangle\langle\psi_{\text{final}}|$ . The fact that in the typical scenario  $|\alpha_{g_{\text{std}}}|^2$  is not necessarily 1 implies that in general,  $\varepsilon_{\text{eff}} \neq 1$ , and the mea-

surement outcomes are correlated. However,  $\varepsilon_{\text{eff}} \approx 1$  and

$$p_{\text{success all}}^q(1, M, N) \approx 1. \quad (29)$$

for  $N \gg 1$ . A more precise lower bound on  $p_{\text{success all}}^q(1, M, N)$  can be obtained via Eqs. (33) and (30); this implies the result of Eq. (29) for  $N \gg 1$ .

The absence of techniques for computing or approximating  $p_{\text{success all}}^q(\varepsilon, M, N)$  for more than a single output bit motivates the search for bounds on this quantity or, equivalently and more conveniently,  $p_{\text{fail all}}^q(\varepsilon, M, N)$ . The strategy is to bound this in terms of the *single bit failure probability*, i.e. the marginal probability with which the algorithm fails to identify a given single qubit correctly regardless of the output for the other bits. Assuming, without loss of generality that the marked item is located at  $s = 11 \dots 11$ , Eq. (2) implies that the joint probability distribution is symmetrical under interchanges of qubits. Thus the single bit failure probability is identical for all qubits and will be denoted  $p_{\text{fail one}}^q(\varepsilon, M, N)$ . The bounding relationship, central to the rest of this article, between the single bit failure probability and the probability with which the algorithm fails is

$$p_{\text{fail one}}^q \leq p_{\text{fail all}}^q \leq n p_{\text{fail one}}^q \quad (30)$$

where the arguments,  $(\varepsilon, M, N)$ , of the probabilities are identical and have been omitted for convenience (see appendix B for a proof).



Following methods established for single bit output algorithms [20],

$$p_{\text{fail one}}^q(\varepsilon, M, N) = \frac{1}{2} \sum_{k=M_{\min}}^M \binom{M}{k} \left(\frac{1-\varepsilon_{\text{eff}}}{2}\right)^k \left(\frac{1+\varepsilon_{\text{eff}}}{2}\right)^{M-k} + \frac{1}{2} \sum_{k=M-M_{\min}+1}^M \binom{M}{k} \left(\frac{1-\varepsilon_{\text{eff}}}{2}\right)^k \left(\frac{1+\varepsilon_{\text{eff}}}{2}\right)^{M-k} \quad (31)$$

where  $\varepsilon_{\text{eff}}$  is given by Eq. (22) and

$$M_{\min} := \left\lceil \frac{M+1}{2} \right\rceil. \quad (32)$$

The single bit failure probability is a monotonically decreasing function of  $\varepsilon$  [20] with  $p_{\text{fail one}}^q(0, M, N) = 1/2$  and

$$p_{\text{fail one}}^q(1, M, N) \leq \sqrt{\frac{2}{\pi M}} \frac{2(N-1)}{2N-3} \left(\frac{2}{N-1}\right)^{M/2} \quad (33)$$

whenever  $M \gg 1$  (see appendix C). Thus for  $N \gg 1$ ,  $p_{\text{fail one}}^q(1, M, N) \approx 0$ . For fixed polarization  $\varepsilon$  and odd  $M$ ,  $p_{\text{fail one}}^q(\varepsilon, M, N) = p_{\text{fail one}}^q(\varepsilon, M+1, N)$  and  $p_{\text{fail one}}^q(\varepsilon, M, N) \geq p_{\text{fail one}}^q(\varepsilon, M+2, N)$  with equality only if  $\varepsilon_{\text{eff}} = 0, 1$  [20].

The exact expression of Eq. (31) is not conducive to rapid numerical evaluation and is less useful for assessing the behavior of  $p_{\text{fail one}}^q(\varepsilon, M, N)$  than certain alternatives. It can easily be verified by repeated integration by parts that

$$\sum_{k=m}^n \binom{n}{k} p^k (1-p)^{n-k} = I_p(m, n-m+1) \quad (34)$$

where

$$I_p(x, y) := \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)} \int_0^p t^{x-1} (1-t)^{y-1} \quad (35)$$

is the incomplete beta function. Applied to Eq. (31) this yields an alternative exact expression,

$$p_{\text{fail one}}^q(\varepsilon, M, N) = \frac{1}{2} [I_p(M_{\min}, M - M_{\min} + 1) + I_p(M - M_{\min} + 1, M_{\min})] \quad (36)$$

where  $p = (1-\varepsilon)/2$ . For  $M \gg 1$ , Eqs. (35) and (36) can be combined (see appendix D) to give the approximation

$$p_{\text{fail one}}^q(\varepsilon, M, N) \approx \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \int_0^{\varepsilon\sqrt{M}} e^{-t^2/2} dt. \quad (37)$$

This is closely related to the standard error function and is more suitable for numerical computation than the exact expression of Eq. (31).

## A. Statistical comparison of the ensemble quantum and classical sequential search algorithms

The ensemble quantum algorithm is probabilistic and it is reasonable to compare it in terms of success probabilities to a probabilistic classical sequential search. A deterministic classical sequential search proceeds by evaluating the oracle,  $f(x)$ , at distinct database locations  $x$ , terminating whenever  $f(x) = 1$ . We consider a probabilistic sequential search in which the oracle is evaluated at  $Q$  distinct randomly chosen database locations. The probability with which this fails to correctly identify the marked item is

$$p_{\text{fail all}}^c(Q, N) = 1 - \frac{Q}{N} \quad (38)$$

provided that  $Q \leq N-2$  since a search with  $N-1$  queries will yield the marked item's location with certainty. We shall compare the two types of searches by requiring that each use *the same resources measured in terms of the aggregate number of oracle invocations* (we briefly discuss other possibilities later). Thus we regard an ensemble implementation invoking  $q$  oracle invocations, each on  $M$  ensemble members, as using equivalent resources to a classical sequential with  $Q = qM$  oracle queries. The central issue in this article is to compare the two types of searches provided that  $Q = qM$  and to determine conditions under which the ensemble implementation of the quantum algorithm outperforms the probabilistic classical sequential search, i.e. such that

$$p_{\text{fail all}}^q(\varepsilon, M, N) < p_{\text{fail all}}^c(qM, N). \quad (39)$$

The *critical polarization*  $\varepsilon_c$ , which demarcates the boundary between the two types of algorithm, is defined implicitly via

$$p_{\text{fail all}}^q(\varepsilon_c, M, N) = p_{\text{fail all}}^c(qM, N). \quad (40)$$

In general  $\varepsilon_c$  can be expected to depend on  $M$  and  $N$  and will be denoted  $\varepsilon_c(M, N)$ . Eqs. (28), (29) and (40) imply that a critical polarization in the range  $0 \leq \varepsilon_c \leq 1$  can only exist provided that

$$1 \leq qM \leq N. \quad (41)$$

The left bound is trivially satisfied while the right bound implies an upper limit on the ensemble size,

$$M_{\max} = \left\lfloor \frac{N}{q} \right\rfloor. \quad (42)$$

For  $M > M_{\max}$ ,  $p_{\text{fail all}}^c(qM, N) = 0$  and the classical algorithm always outperforms the quantum algorithm, except in some (but not all) situations for which  $\varepsilon = 1$ . For  $M < M_{\max}$  the critical polarization exists since  $p_{\text{fail all}}^q(\varepsilon, M, N)$  is a continuous function of  $\varepsilon$  and the classical failure probability lies within the bounds for the quantum failure probability.

The difficulties in calculating  $p_{\text{success all}}^q(\varepsilon, M, N)$  are equivalent to those in calculating  $p_{\text{fail all}}^q(\varepsilon, M, N)$  and we resort to calculating bounds, originating from the inequality of Eq. (30), for the critical polarization. We shall show that

$$\varepsilon_{\text{c necc}}(M, N) \leq \varepsilon_{\text{c}}(M, N) \leq \varepsilon_{\text{c suff}}(M, N) \quad (43)$$

where the *necessary critical polarization*,  $\varepsilon_{\text{c necc}}$ , satisfies

$$p_{\text{fail one}}^q(\varepsilon_{\text{c necc}}, M, N) = p_{\text{fail all}}^c(qM, N). \quad (44)$$

and the *sufficient critical polarization*,  $\varepsilon_{\text{c suff}}$ , satisfies

$$p_{\text{fail one}}^q(\varepsilon_{\text{c suff}}, M, N) = \frac{1}{n} p_{\text{fail all}}^c(qM, N). \quad (45)$$

Each of these depend on  $M$  and  $N$  and they will sometimes be denoted  $\varepsilon_{\text{c necc}}(M, N)$  and  $\varepsilon_{\text{c suff}}(M, N)$ . As the necessary and sufficient critical polarizations are defined via the single bit failure probability, existing techniques [20] can be used to demonstrate their properties and calculate them.

The existence of the necessary critical polarization follows from the fact [20] that  $p_{\text{fail one}}^q(\varepsilon, M, N)$  is a monotonically decreasing function of  $\varepsilon$  ranging from 0 to  $1/2$ . Thus a unique value of the necessary critical polarization exists whenever  $\lceil N/2q \rceil \leq M \leq M_{\text{max}}$ . Note that in the typical scenario this implies  $M_{\text{max}}/2 \leq M \leq M_{\text{max}}$ . For  $M \leq \lceil N/2q \rceil$ , the classical sequential search fails with probability larger than  $1/2$  and thus the ensemble quantum algorithm always outperforms it, giving  $\varepsilon_{\text{c necc}}(M, N) = 0$ . Similarly, the sufficient critical polarization exists provided that  $(1 - n/2)N/q \leq M \leq M_{\text{max}}$ . The left hand inequality is satisfied whenever  $n \geq 2$  as it does in the typical scenario. Here a unique sufficient polarization exists when  $0 \leq M \leq M_{\text{max}}$ .

The term “necessary” stems from considering  $\varepsilon < \varepsilon_{\text{c necc}}$ , in which case the monotonic behavior of  $p_{\text{fail one}}^q(\varepsilon, M, N)$  with respect to  $\varepsilon$  and Eqs (30) and (44) imply

$$p_{\text{fail all}}^q(\varepsilon, M, N) > p_{\text{fail all}}^c(qM, N); \quad (46)$$

here the classical algorithm succeeds with greater probability than the quantum ensemble algorithm. In this sense,  $\varepsilon_{\text{c necc}}$  establishes a lower bound on  $\varepsilon_{\text{c}}$ .

The term “sufficient” arises by considering,  $\varepsilon > \varepsilon_{\text{c suff}}$ , for which a similar argument implies

$$p_{\text{fail all}}^q(\varepsilon, M, N) < p_{\text{fail all}}^c(qM, N); \quad (47)$$

here the quantum ensemble algorithm succeeds with greater probability than the classical sequential search using the same resources. In this regard  $\varepsilon_{\text{c suff}}$  is an upper bound on  $\varepsilon_{\text{c}}$ . This establishes Eq. (43).

The task now becomes one of computing the two bounding critical polarizations. Eqs. (38) and (44) with

$Q = qM$  imply that

$$\begin{aligned} p_{\text{fail one}}^q(\varepsilon_{\text{c necc}}, M, N) &= 1 - \frac{qM}{N} \\ &\approx 1 - \frac{\pi M}{4\sqrt{N}}. \end{aligned} \quad (48)$$

For fixed  $N$ , Eq. (48) can be solved numerically for  $\varepsilon_{\text{c necc}}(M, N)$ . Data, computed via Eq. (36) using MATHEMATICA’s implementation of the incomplete beta function, is plotted in Fig. 1.

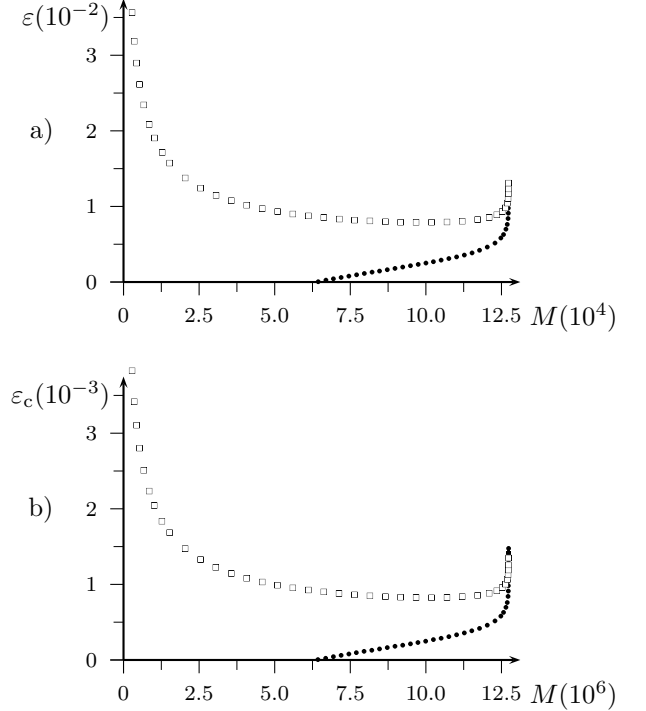


FIG. 1: Critical polarization for a)  $N = 10^{10}$  and b)  $N = 10^{14}$ . Open squares indicate sufficient polarization and solid dots necessary polarization. Note the distinct horizontal and vertical scales. The rightmost point in each plot corresponds to  $M = M_{\text{max}}$  for the relevant value of  $N$ . The critical polarization,  $\varepsilon_{\text{c}}$ , lies between the two curves.

In general it can be shown (see appendix E) that  $\varepsilon_{\text{c necc}}(M, N)$  is a monotonically increasing function of  $M$  with  $\varepsilon_{\text{c necc}} \rightarrow 1$  as  $M \rightarrow M_{\text{max}}$ . A more important consideration is the scaling behavior of  $\varepsilon_{\text{c necc}}(M, N)$  with respect to  $N$  while fixing  $M$  suitably. The similarity in the data generated for the two values of  $N$  illustrated in Fig. 1 suggests an approach based on  $\varepsilon_{\text{c necc}}$  as a function  $M/M_{\text{max}}$  for different values of  $N$ . In fact, it is easily shown that

$$p_{\text{fail all}}^c(qM, N) \approx 1 - \frac{M}{M_{\text{max}}} \quad (49)$$

where  $M_{\text{max}}$  depends on  $q$  and  $N$  via Eq. (42). Thus the necessary critical polarization is approximately determined via

$$p_{\text{fail one}}^q(\varepsilon_{\text{c necc}}, M, N) = 1 - \frac{M}{M_{\text{max}}}. \quad (50)$$

For  $N \gg 1$ , Eqs. (22) and (31) imply that the left hand side is independent of  $N$  and  $q$  and these only enter into the problem via  $M_{\max}$ . Our strategy for determining the scaling behavior of  $\varepsilon_{\text{c necc}}$  with respect to  $N$  is to consider fixed  $M/M_{\max}$  as  $N$  varies. This is equivalent to maintaining an ensemble size needed to sustain a fixed failure probability for the classical algorithm. In the typical scenario  $M_{\max} \approx \pi\sqrt{N}/4$  and upon multiplying  $N$  by a factor  $\gamma$ , the ratio  $M/M_{\max}$  can only be held constant by multiplying  $M$  by a factor of  $\sqrt{\gamma}$ . More explicitly, for any  $\gamma > 0$ , Eq. (48) implies that

$$\begin{aligned} p_{\text{fail one}}^q(\varepsilon_{\text{c necc}}(\sqrt{\gamma}M, \gamma N), \sqrt{\gamma}M, \gamma N) \\ = p_{\text{fail one}}^q(\varepsilon_{\text{c necc}}(M, N), M, N) \end{aligned}$$

and the approximation of Eq. (37) implies

$$\varepsilon_{\text{c necc}}(\sqrt{\gamma}M, \gamma N) \approx \gamma^{-1/4} \varepsilon_{\text{c necc}}(M, N) \quad (51)$$

or equivalently

$$\varepsilon_{\text{c necc}}(M, \gamma N) \approx \gamma^{-1/4} \varepsilon_{\text{c necc}}(M/\sqrt{\gamma}, N). \quad (52)$$

This relationship is evident in the numerical data; the plots of Fig. 1 illustrate an instance where  $\gamma = 10^4$ .

The sufficient polarization satisfies

$$\begin{aligned} p_{\text{fail one}}^q(\varepsilon_{\text{c suff}}, M, N) &= \frac{1}{n} \left( 1 - \frac{qM}{N} \right) \\ &\approx \frac{1}{n} \left( 1 - \frac{M}{M_{\max}} \right) \end{aligned} \quad (53)$$

Data generated via numerical solution of Eq. (53), using the incomplete beta function representation of Eq. (36), is illustrated in Fig. 1. A notable feature of the data is that for much of the region in which  $M_{\max}/2 \leq M \leq M_{\max}$ , the necessary and sufficient critical polarizations are within an order of magnitude of each other; if all that is of interest is the typical order of magnitude of  $\varepsilon_{\text{c}}$  for a given  $N$  then it suffices to compute either  $\varepsilon_{\text{c necc}}$  or  $\varepsilon_{\text{c suff}}$  for  $M$  a moderately large fraction of  $M_{\max}$ . By this argument one can attain an order of magnitude estimate of the critical polarization, which lies between the bounding critical polarizations. Additionally, for much of the range  $0 \leq M \leq M_{\max}$ , the numerical data indicates that the sufficient polarization is roughly constant.

In the limiting case  $\varepsilon_{\text{c suff}} \rightarrow 1$  as  $M \rightarrow M_{\max}$ . Also as  $M \rightarrow 1$ ,  $p_{\text{fail one}}^q(\varepsilon_{\text{c suff}}, M, N) \rightarrow (1 - \varepsilon)/2$  and thus Eq. (53) implies that

$$\varepsilon_{\text{c suff}} \rightarrow 1 - \frac{2}{n} \left( 1 - \frac{\pi}{4\sqrt{N}} \right) \approx 1 \quad (54)$$

in the typical scenario. The scaling of the sufficient critical polarization is more complicated than that of the necessary critical polarization as a result of the additional factor of  $1/n$  in Eq. (53). A similar line of reasoning to that for the necessary polarization gives

$$\begin{aligned} p_{\text{fail one}}^q(\varepsilon_{\text{c suff}}(\sqrt{\gamma}M, \gamma N), \sqrt{\gamma}M, \gamma N) \\ = \frac{\lceil \log_2 N \rceil}{\lceil \log_2(\gamma N) \rceil} p_{\text{fail one}}^q(\varepsilon_{\text{c suff}}(M, N), M, N) \end{aligned} \quad (55)$$

Here the terms involving logarithms prohibit immediate use of Eq. (37) as was done for the necessary polarizations. However, when  $\lceil \log_2 \gamma \rceil \ll \lceil \log_2 N \rceil$ ,

$$\varepsilon_{\text{c suff}}(\sqrt{\gamma}M, \gamma N) \approx \gamma^{-1/4} \varepsilon_{\text{c suff}}(M, N), \quad (56)$$

since the factor involving logarithms on the right of Eq. (55) is approximately unity. Again this is evident for the case of  $\gamma = 10^4$  as illustrated in Fig. 1.

An additional illustration of the scaling properties of the necessary and sufficient critical polarizations can be attained by computing these at values of  $M$  for which the classical algorithm returns a fixed success probability rate as  $N$  varies. In order to succeed with probability,  $0 \leq p \leq 1$ , the classical algorithm requires  $Q = \lceil pN \rceil$  oracle queries. The ensemble quantum algorithm uses comparable resources for  $M = \lceil Q/q \rceil \approx 4p/\pi\sqrt{N} \approx pM_{\max}$  ensemble members. Necessary and sufficient critical polarizations can be computed using  $N$  and the value of  $M$  computed from it as just described. This method of fixing  $M$  is the same as that resulting in Eqs. (51) and (56). Dropping the first argument, which is computed from the second and  $p$ , Eq. (51) is equivalent to

$$\varepsilon_{\text{c necc}}(N) = \left( \frac{N}{N_0} \right)^{-1/4} \varepsilon_{\text{c necc}}(N_0) \quad (57)$$

where  $N_0$  is any fixed database size that is sufficiently large for the approximations to be valid. A similar conclusion applies to the sufficient polarization. For  $p = 0.90$ , numerically generated data is illustrated, using logarithmic scaling, in Fig. 2.

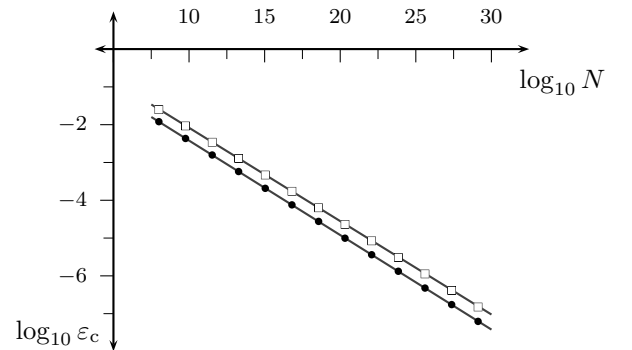


FIG. 2: Critical polarization for success probability of 0.90. Open squares indicate sufficient critical polarization and solid dots necessary critical polarization. The straight lines are obtained by least squares fits. Similar plots are obtained for other values of  $p$ , the primary difference being that the lines are shifted vertically.

Least squares linear fits for these data give

$$\begin{aligned} \log_{10} \varepsilon_{\text{c necc}} &= 0.078 - 0.250 \log_{10} N \\ \log_{10} \varepsilon_{\text{c suff}} &= 0.387 - 0.247 \log_{10} N \end{aligned}$$

which are consistent with Eqs. (57); the discrepancy in the coefficient of  $\log_{10} N$  for  $\varepsilon_{\text{c suff}}$  probably arising from



the factor of  $1/n$  in the definition of the sufficient polarization.

We now briefly consider an alternative method of comparing the two types of algorithm. Suppose that the ensemble algorithm uses fewer resources than the probabilistic sequential search in the sense that  $(qM)^\alpha = Q$  where  $\alpha > 1$ . The fact that the classical succeeds with certainty when  $Q \geq N - 1$  means that, in the typical scenario, the ensemble implementation can only outperform its classical competitor when  $\alpha < 2$ . It is easy to show that for the necessary polarization,  $0 \leq M \leq M_{\max}$ , where now  $M_{\max} = \lfloor N^{(1/\alpha)}/q \rfloor$  and for the typical scenario this leaves  $M_{\max} \approx 4N^{(1/\alpha-1/2)}/\pi$ . The scaling argument can be repeated by considering multiplication of  $N$  by a factor  $\gamma > 0$ . A constant classical failure probability is then maintained by considering  $M$  multiplied by a factor of  $\gamma^{(1/\alpha-1/2)}$ . Repeating the argument involving the single bit failure probability for the ensemble quantum algorithm ultimately yields

$$\varepsilon_{\text{c necc}}(N) = \left( \frac{N}{N_0} \right)^{[1/4-1/(2\alpha)]} \varepsilon_{\text{c necc}}(N_0) \quad (58)$$

with a similar inference for the sufficient polarization.

Current room-temperature, solution state NMR using pseudopure preparation schemes [6, 24, 25] attains typical polarizations of  $\varepsilon \sim 10^{-5}$ . The numerical data plotted in Fig. 1 and the scaling of Eq. (51) and (56) suggest that these standard realizations would only gain an advantage for databases of size  $N \gtrsim 10^{22}$ . In this respect, the requirements on the ensemble for an advantageous realization of the Grover search algorithm are less demanding than those for the Deutsch-Jozsa algorithm [20], where it was found that  $\varepsilon \geq 0.866$  was necessary in the best case, regardless of the problem size. It should be noted that the standard number of oracle queries for such database sizes is of the order of  $10^{11}$ , which forms a lower bound for the number of pulses in sequence of gates; sufficient coherent control over such a number of pulses is evidently beyond current NMR experimental capabilities. Nevertheless, compared to the situation of the ensemble Deutsch-Jozsa algorithm, our results are promising since an increase of one order of magnitude in polarization will result advantages for searches on a database whose size decreases by four orders of magnitude. In fact, for a dramatic increase in polarization such as that produced by newer schemes [26, 27] producing a large initial polarization of  $\varepsilon \sim 0.9$  would be sufficient for modest database sizes where  $N \lesssim 10^6$ . Also, to date, all solution state NMR realizations of Grover's algorithm [11, 16, 28] have involved databases for which  $N \leq 8$  and ensemble sizes of  $M \sim 10^{20}$ . Thus  $M \gg M_{\max}$  in these cases and, according to our criteria, the ensemble quantum algorithm uses greater resources than a classical sequential search which will locate the marked item with certainty.

## V. CONCLUSION

In conclusion, we have extended methods for statistical comparisons of single bit output ensemble quantum algorithms to classical probabilistic competitors to the Grover search algorithm. Our results indicate that the polarization required for the ensemble quantum algorithm to outperform known classical competitors scales as  $N^{-1/4}$  where  $N$  is the database size and the ensemble size is scaled so as to fix the failure probability of the classical algorithm. In absolute terms, the required polarization is modest compared to that required for an advantageous ensemble realization of the Deutsch-Jozsa algorithm.

## Acknowledgments

Much of this work was carried out while the authors were in the Department of Physics and Astronomy at Bucknell University, Lewisburg, PA. We would like to thank that institution for the use of its resources and facilities as well as for support and encouragement from our colleagues there.

## APPENDIX A: SUCCESS PROBABILITY FOR TWO QUBITS

We aim to compute the probability with which qubit 1 returns the correct value  $k_1$  times and qubit 2 returns the correct value  $k_2$  times. Consider the joint measurement outcomes on each ensemble member; denote a typical outcome by  $x_2 x_1$  where the subscript indicates the qubit number. Let  $l_0$  denote the number of times that 00 occurs,  $l_1$  the number of times that 01 occurs,  $l_2$  the number of times that 10 occurs, etc . . . . The probability with which this occurs is

$$(l_0, l_1, l_2, l_3)! \left( \frac{1-\varepsilon}{4} \right)^{l_0} \left( \frac{1-\varepsilon}{4} \right)^{l_1} \left( \frac{1-\varepsilon}{4} \right)^{l_2} \left( \frac{1+3\varepsilon}{4} \right)^{l_3} \quad (A1)$$

where the multinomial symbol is

$$(l_0, l_1, l_2, l_3)! := \frac{(l_0 + l_1 + l_2 + l_3)!}{l_0! l_1! l_2! l_3!}. \quad (A2)$$

Given values for  $k_1$  and  $k_2$ , various combinations of  $l_0, l_1, l_2, l_3$  are possible; all are subject to the following constraints:

$$l_0, l_1, l_2, l_3 \geq 0 \quad (A3a)$$

$$l_0 + l_1 + l_2 + l_3 = M \quad (A3b)$$

$$l_1 + l_3 = k_1 \quad (A3c)$$

$$l_2 + l_3 = k_2. \quad (A3d)$$

Thus

$$l_2 = k_1 - l_3 \quad (\text{A4a})$$

$$l_1 = k_2 - l_3 \quad (\text{A4b})$$

$$l_0 = M - k_1 - k_2 + l_3 \quad (\text{A4c})$$

and together with Eq. (A3a) these give

$$l_3 \geq k_1 + k_2 - M \quad (\text{A5a})$$

$$k_1 \geq l_3 \quad (\text{A5b})$$

$$k_2 \geq l_3 \quad (\text{A5c})$$

$$l_3 \geq 0 \quad (\text{A5d})$$

Eqs. (A5b) and (A5c) are equivalent to

$$l_3 \leq \min(k_1, k_2). \quad (\text{A6})$$

while Eqs. (A5a) and (A5d) are equivalent to

$$l_3 \geq \max(0, k_1 + k_2 - M). \quad (\text{A7})$$

Any value of  $l_3$  such that  $\max(0, k_1 + k_2 - M) \leq l_3 \leq \min(k_1, k_2)$  together with Eqs. (A4) yield  $k_1$  and  $k_2$ . Thus

$$\begin{aligned} \Pr(k_1, k_2) &= \sum_{l_3=\max(0, k_1+k_2-M)}^{\min(k_1, k_2)} (M - k_1 - k_2 + l_3, k_1 - l_3, k_2 - l_3, l_3)! \\ &\quad \times \left(\frac{1-\varepsilon}{4}\right)^{M-k_1-k_2+l_3} \left(\frac{1-\varepsilon}{4}\right)^{k_1-l_3} \left(\frac{1-\varepsilon}{4}\right)^{k_2-l_3} \left(\frac{1+3\varepsilon}{4}\right)^{l_3} \\ &= \frac{1}{4^M} \sum_{l=\max(0, k_1+k_2-M)}^{\min(k_1, k_2)} (M - k_1 - k_2 + l, k_1 - l, k_2 - l, l)! (1-\varepsilon)^{M-l} (1+3\varepsilon)^l. \end{aligned} \quad (\text{A8})$$

## APPENDIX B: UPPER BOUND ON FAILURE PROBABILITY

The inequality

$$p_{\text{fail all}}^q(\varepsilon, M, N) \leq n p_{\text{fail one}}^q(\varepsilon, M, N) \quad (\text{B1})$$

is clearly true for  $n = 1$ . We adopt the following notation for the  $n$  qubit case:  $\Pr(s_n, \dots, s_3, f_2, s_1)$  is the probability that we succeed in identifying bit 1 correctly, fail on bit 2, succeed on bit 3 and so on. For  $n = 2$ ,

$$\begin{aligned} p_{\text{fail all}}^q(\varepsilon, M, N) &= \Pr(f_2, f_1) + \Pr(f_2, s_1) + \Pr(s_2, f_1) \\ &= \Pr(f_2, f_1) + \Pr(f_2, s_1) + \Pr(s_2, f_1) \\ &\quad + \Pr(f_2, f_1) - \Pr(f_2, f_1). \end{aligned}$$

However, for bit 2,

$$p_{\text{fail one}}^q(\varepsilon, M, N) = \Pr(f_2, f_1) + \Pr(f_2, s_1) \quad (\text{B2})$$

with a similar result for bit 2. Thus

$$\begin{aligned} p_{\text{fail all}}^q(\varepsilon, M, N) &= 2 p_{\text{fail one}}^q(\varepsilon, M, N) - \Pr(f_2, f_1) \\ &\leq 2 p_{\text{fail one}}^q(\varepsilon, M, N) \end{aligned}$$

since  $\Pr(f_2, f_1) \geq 0$ . The general result follows by induction on the number of qubits. For  $n$  qubits,

$$\begin{aligned}
p_{\text{fail all}}^q(\varepsilon, M, N) &= \Pr(f_n, \text{any outcome on first } n-1) + \Pr(s_n, \text{fail in some way on first } n-1) \\
&= p_{\text{fail one}}^q(\varepsilon, M, N) + \Pr(s_n, \text{fail in some way on first } n-1) \\
&= p_{\text{fail one}}^q(\varepsilon, M, N) + \Pr(s_n, \text{fail in some way on first } n-1) \\
&\quad + \Pr(f_n, \text{fail in some way on first } n-1) - \Pr(f_n, \text{fail in some way on first } n-1) \\
&\leq p_{\text{fail one}}^q(\varepsilon, M, N) + \Pr(s_n, \text{fail in some way on first } n-1) \\
&\quad + \Pr(f_n, \text{fail in some way on first } n-1) \\
&\leq p_{\text{fail one}}^q(\varepsilon, M, N) + (n-1)p_{\text{fail one}}^q(\varepsilon, M, N) \\
&\leq np_{\text{fail one}}^q(\varepsilon, M, N)
\end{aligned} \tag{B3}$$


---

which proves the result.

---

### APPENDIX C: PERFECT POLARIZATION BOUND FOR SINGLE BIT FAILURE PROBABILITY

The single bit failure probability is

$$\begin{aligned}
p_{\text{fail one}}^q(\varepsilon, M, N) &= \frac{1}{2} \sum_{k=M_{\min}}^M \binom{M}{k} \left(\frac{1-\varepsilon_{\text{eff}}}{2}\right)^k \left(\frac{1+\varepsilon_{\text{eff}}}{2}\right)^{M-k} \\
&\quad + \frac{1}{2} \sum_{k=M-M_{\min}+1}^M \binom{M}{k} \left(\frac{1-\varepsilon_{\text{eff}}}{2}\right)^k \left(\frac{1+\varepsilon_{\text{eff}}}{2}\right)^{M-k}.
\end{aligned} \tag{C1}$$


---

For  $\varepsilon = 1$  and  $q = q_{\text{std}}$ ,

$$1 - \frac{1}{N} \leq |\alpha_q|^2 \leq 1 \tag{C2}$$

giving

$$1 - \frac{1}{N-1} \leq \varepsilon_{\text{eff}} \leq 1. \tag{C3}$$

Thus

$$\begin{aligned}
\frac{1-\varepsilon_{\text{eff}}}{2} &\leq \frac{1}{2(N-1)} \\
\frac{1+\varepsilon_{\text{eff}}}{2} &\leq 1
\end{aligned}$$

which implies that the first term in Eq (C1) satisfies

$$\begin{aligned}
\frac{1}{2} \sum_{k=M_{\min}}^M \binom{M}{k} \left(\frac{1-\varepsilon_{\text{eff}}}{2}\right)^k \left(\frac{1+\varepsilon_{\text{eff}}}{2}\right)^{M-k} &\leq \frac{1}{2} \sum_{k=M_{\min}}^M \frac{1}{2^k} \frac{1}{(N-1)^k} \\
&= \frac{1}{2} \binom{M}{M_{\min}} \frac{1}{[2(N-1)]^{M_{\min}}} \\
&\quad \times \left\{ 1 + \frac{1}{2(N-1)} \binom{M}{M_{\min}+1} / \binom{M}{M_{\min}} + \dots \right\} \\
&\leq \frac{1}{2} \binom{M}{M_{\min}} \frac{1}{[2(N-1)]^{M_{\min}}} \left\{ 1 + \frac{1}{2(N-1)} + \dots \right\} \\
&\leq \frac{1}{2} \binom{M}{M_{\min}} \frac{1}{[2(N-1)]^{M_{\min}}} \frac{2(N-1)}{2N-3}
\end{aligned} \tag{C4}$$


---

since

$$\sum_{k=0}^{\infty} \frac{1}{x^k} = \frac{x}{x-1}. \tag{C5}$$

Stirling's approximation gives [20]

$$\binom{M}{M_{\min}} \approx 2^M \sqrt{\frac{2}{\pi M}} \tag{C6}$$

for  $M \gg 1$  and thus the first term in Eq (C1) is bounded where  
by

$$\begin{aligned} & \frac{1}{2} \sum_{k=M_{\min}}^M \binom{M}{k} \left( \frac{1-\varepsilon_{\text{eff}}}{2} \right)^k \left( \frac{1+\varepsilon_{\text{eff}}}{2} \right)^{M-k} \\ & \leq \frac{1}{2} \sqrt{\frac{2}{\pi M}} \frac{2(N-1)}{2N-3} \left( \frac{2}{N-1} \right)^{M/2}. \end{aligned} \quad (\text{C7})$$

A similar argument applies to the second term in Eq (C1), giving the same bound. Thus

$$p_{\text{fail one}}^{\text{q}}(1, M, N) \leq \sqrt{\frac{2}{\pi M}} \frac{2(N-1)}{2N-3} \left( \frac{2}{N-1} \right)^{M/2}. \quad p = (1-\varepsilon)/2, \text{ and} \quad (\text{C8})$$

#### APPENDIX D: GAUSSIAN APPROXIMATION TO THE QUANTUM FAILURE PROBABILITY

The single bit failure probability can be expressed exactly in terms of the incomplete beta function

$$\begin{aligned} p_{\text{fail one}}^{\text{q}}(\varepsilon, M, N) = & \frac{1}{2} [I_p(M_{\min}, M - M_{\min} + 1) \\ & + I_p(M - M_{\min} + 1, M_{\min})] \end{aligned} \quad (\text{D1})$$

Substituting from Eq. (D2) into Eq. (D1) gives

$$M_{\min} = \left\lceil \frac{M+1}{2} \right\rceil. \quad (\text{D3})$$

$$\begin{aligned} p_{\text{fail one}}^{\text{q}}(\varepsilon, M, N) = & \frac{1}{2} \frac{\Gamma(M+1)}{\Gamma(M_{\min})\Gamma(M - M_{\min} + 1)} \int_0^p [t^{M_{\min}-1}(1-t)^{M-M_{\min}} \\ & + t^{M-M_{\min}}(1-t)^{M_{\min}-1}] dt. \end{aligned} \quad (\text{D4})$$

Redefining the variable of integration via  $t' := 1 - 2t$  gives

$$\begin{aligned} p_{\text{fail one}}^{\text{q}}(\varepsilon, M, N) = & \frac{1}{2^{M+1}} \frac{\Gamma(M+1)}{\Gamma(M_{\min})\Gamma(M - M_{\min} + 1)} \int_{\varepsilon}^1 [(1-t')^{M_{\min}-1}(1+t')^{M-M_{\min}} \\ & + (1-t')^{M-M_{\min}}(1+t')^{M_{\min}-1}] dt' \\ = & \frac{1}{2^{M+1}} \frac{\Gamma(M+1)}{\Gamma(M_{\min})\Gamma(M - M_{\min} + 1)} \left\{ \int_0^1 [\dots] dt' - \int_0^{\varepsilon} [\dots] dt' \right\}. \end{aligned} \quad (\text{D5})$$

However,

$$p_{\text{fail one}}^{\text{q}}(0, M, N) = \frac{1}{2^{M+1}} \frac{\Gamma(M+1)}{\Gamma(M_{\min})\Gamma(M - M_{\min} + 1)} \int_0^1 [\dots] dt' = \frac{1}{2} \quad (\text{D6})$$

and thus

$$\begin{aligned} p_{\text{fail one}}^{\text{q}}(\varepsilon, M, N) = & \frac{1}{2} - \frac{1}{2^{M+1}} \frac{\Gamma(M+1)}{\Gamma(M_{\min})\Gamma(M - M_{\min} + 1)} \int_0^{\varepsilon} [(1-t')^{M_{\min}-1}(1+t')^{M-M_{\min}} \\ & + (1-t')^{M-M_{\min}}(1+t')^{M_{\min}-1}] dt'. \end{aligned} \quad (\text{D7})$$

The gamma function satisfies  $\Gamma(n+1) = n!$  for integral  $n$ , giving

$$\frac{\Gamma(M+1)}{\Gamma(M_{\min})\Gamma(M - M_{\min} + 1)} = \frac{M!}{(M - M_{\min})!(M_{\min} - 1)!}. \quad (\text{D8})$$

For  $M \gg 1$ ,  $M_{\min} \approx M/2$  and Stirling's approximation gives [20]

$$\frac{\Gamma(M+1)}{\Gamma(M_{\min})\Gamma(M - M_{\min} + 1)} \approx 2^M \sqrt{\frac{M}{2\pi}}. \quad (\text{D9})$$

Considering the expression of Eq. (D7), each of the exponents within the integral are approximately  $M/2$ . Thus

$$p_{\text{fail one}}^q(\varepsilon, M, N) \cong \frac{1}{2} - \sqrt{\frac{M}{2\pi}} \int_0^\varepsilon (1-t'^2)^{M/2} dt'. \quad (\text{D10})$$

With  $t'' := \sqrt{M}t'$ , Eq. (D10) implies

$$p_{\text{fail one}}^q(\varepsilon, M, N) \cong \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \int_0^{\varepsilon\sqrt{M}} \left(1 - \frac{t''^2}{M}\right)^{M/2} dt'' \quad (\text{D11})$$

and for  $M \gg 1$  the integrand approximates an exponential function, giving

$$p_{\text{fail one}}^q(\varepsilon, M, N) \cong \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \int_0^{\varepsilon\sqrt{M}} e^{-t^2/2} dt. \quad (\text{D12})$$

## APPENDIX E: NECESSARY CRITICAL POLARIZATION VS $M$

We show that  $\varepsilon_{c \text{ necc}}(M, N)$  is a monotonically increasing function of  $M$ . In general  $\varepsilon_{c \text{ necc}}$  satisfies

$$p_{\text{fail one}}^q(\varepsilon_{c \text{ necc}}(M, N), q, N)M = 1 - \frac{qM}{N}. \quad (\text{E1})$$

For  $M$  odd and any  $\varepsilon$  it can be shown [20] that  $p_{\text{fail one}}^q(\varepsilon, M, N) = p_{\text{fail one}}^q(\varepsilon, M+1, N)$ . The right side of Eq. (E1) decreases as  $M$  increases. Thus  $p_{\text{fail one}}^q(\varepsilon_{c \text{ necc}}(M+1, N), M+1, N) < p_{\text{fail one}}^q(\varepsilon_{c \text{ necc}}(M, N), M, N)$  for odd  $M$  and since  $p_{\text{fail one}}^q(\varepsilon, M, N)$  decreases monotonically as  $\varepsilon$  increases, this implies that, for odd  $M$ ,  $\varepsilon_{c \text{ necc}}(M+1, N) > \varepsilon_{c \text{ necc}}(M, N)$ . It remains to consider  $\varepsilon_{c \text{ necc}}(M+2, N)$  versus  $\varepsilon_{c \text{ necc}}(M, N)$  for odd  $M$ . Consider the ratio of the success probabilities,

$$r(\varepsilon, M, N) := \frac{p_{\text{success one}}^q(\varepsilon, M, N)}{p_{\text{success all}}^c(qM, N)} \quad (\text{E2})$$

where  $p_{\text{success all}}^c(qM, N) = qM/N$  is the probability with which the classical sequential search succeeds and  $p_{\text{success one}}^q(\varepsilon, M, N) = 1 - p_{\text{fail one}}^q(\varepsilon, M, N)$ . We shall prove that  $\Delta r(\varepsilon, M, N) := r(\varepsilon, M+2, N) - r(\varepsilon, M, N) < 0$  whenever  $M$  is odd. Then  $r(\varepsilon_{c \text{ necc}}(M, N), M, N) = 1$  implies that  $r(\varepsilon_{c \text{ necc}}(M, N), M+2, N) < r(\varepsilon_{c \text{ necc}}(M, N), M, N) = 1$ . Thus  $\varepsilon_{c \text{ necc}}(M+2, N) > \varepsilon_{c \text{ necc}}(M, N)$ . Now, for  $M$  odd [20],

$$\begin{aligned} p_{\text{success one}}^q(\varepsilon, M, N) &= \sum_{k=\frac{M+1}{2}}^M \binom{M}{k} \left(\frac{1+\varepsilon}{2}\right)^k \left(\frac{1-\varepsilon}{2}\right)^{M-k} \\ &= I_p\left(\frac{M+1}{2}, \frac{M+1}{2}\right) \end{aligned} \quad (\text{E3})$$

where  $p = (1+\varepsilon)/2$ . In these terms

$$\begin{aligned} \Delta r(\varepsilon, M, N) &= \frac{N}{q} \left[ \frac{1}{M+2} I_p\left(\frac{M+3}{2}, \frac{M+3}{2}\right) - \frac{1}{M} I_p\left(\frac{M+1}{2}, \frac{M+1}{2}\right) \right]. \end{aligned} \quad (\text{E4})$$

Then

$$\begin{aligned} I_p\left(\frac{M+3}{2}, \frac{M+3}{2}\right) &= \frac{\Gamma(M+3)}{\Gamma(\frac{M+1}{2}+1)\Gamma(\frac{M+1}{2}+1)} \int_0^p t^{(M+1)/2} (1-t)^{(M+1)/2} dt \\ &= \frac{(M+2)(M+1)}{[(M+1)/2]^2} \frac{\Gamma(M+1)}{\Gamma(\frac{M+1}{2})\Gamma(\frac{M+1}{2})} \int_0^p t^{(M+1)/2} (1-t)^{(M+1)/2} dt \end{aligned} \quad (\text{E5})$$

and thus

$$\begin{aligned} \Delta r(\varepsilon, M, N) &= \frac{N}{q} \frac{\Gamma(M+1)}{[\Gamma(\frac{M+1}{2})]^2} \int_0^p \left\{ \frac{4}{M+1} t^{(M+1)/2} (1-t)^{(M+1)/2} - \frac{1}{M} t^{(M+1)/2} (1-t)^{(M+1)/2} \right\} dt \\ &= \frac{N}{q} \frac{\Gamma(M+1)}{[\Gamma(\frac{M+1}{2})]^2} \int_0^p \left\{ \frac{4}{M+1} t(1-t) - \frac{1}{M} \right\} t^{(M-1)/2} (1-t)^{(M-1)/2} dt. \end{aligned} \quad (\text{E6})$$

For  $0 \leq t \leq 1$  it is easily verified that  $0 \leq t(1-t) \leq 1/4$ . Thus the term in parenthesis within the integral of

Eq. (E6) is negative (except at the single point  $t = 1/2$ )



and  $\Delta r(\varepsilon, M, N) < 0$ .

- 
- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
  - [2] M. L. Bellac, *Short Introduction to Quantum Information and Quantum Computation* (Cambridge University Press, Cambridge, 2006).
  - [3] A. Ekert and R. Jozsa, *Rev. Mod. Phys.* **68**, 733 (1996).
  - [4] A. Steane, *Rep. Prog. Phys.* **61**, 117 (1998).
  - [5] D. G. Cory, M. D. Price, and T. F. Havel, *Physica D* **120**, 82 (1998).
  - [6] I. L. Chuang, N. Gershenfeld, M. G. Kubinec, and D. W. Leung, *Proc. R. Soc. Lond. A* **454**, 447 (1998).
  - [7] D. G. Cory, A. F. Fahmy, and T. F. Havel, *Proc. Nat. Acad. Sci.* **94**, 1634 (1997).
  - [8] N. A. Gershenfeld and I. L. Chuang, *Science* **275**, 350 (1997).
  - [9] R. Marx, A. F. Fahmy, J. M. Myers, W. Bermel, and S. J. Glaser, *Phys. Rev. A* **62**, 012310 (2000).
  - [10] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, R. Cleve, and I. L. Chuang, *Nature* **414**, 883 (2001).
  - [11] L. M. K. Vandersypen, M. Steffen, M. H. Sherwood, C. S. Yannoni, G. Breyta, and I. L. Chuang, *App. Phys. Lett.* **76**, 646 (2000).
  - [12] L. M. K. Vandersypen and I. L. Chuang, *Reviews of Modern Physics* **76**, 1037, (2004).
  - [13] C. Negrevergne, R. Somma, G. Ortiz, E. Knill, and R. Laflamme, *Physical Review A (Atomic, Molecular, and Optical Physics)* **71**, 032344, (2005).
  - [14] C. Negrevergne, T. S. Mahesh, C. A. Ryan, M. Ditty, F. Cyr-Racine, W. Power, N. Boulant, T. Havel, D. G. Cory, and R. Laflamme, *Physical Review Letters* **96**, 170501, (2006).
  - [15] L. J. Schulman and U. Vazirani, *Proc. 31st ACM Symposium on Theory of Computing* p. 322 (1999).
  - [16] I. L. Chuang, N. Gershenfeld, and M. Kubinec, *Phys. Rev. Lett.* **80**, 3408 (1998).
  - [17] E. Knill, I. Chuang, and R. Laflamme, *Phys. Rev. A* **57**, 3348 (1998).
  - [18] D. Collins, *Phys. Rev. A* **65**, 052321 (2002).
  - [19] E. Knill and R. Laflamme, *Phys. Rev. Lett.* **81**, 5672 (1998).
  - [20] B. M. Anderson and D. Collins, *Phys. Rev. A* **72**, 042337 (2005).
  - [21] Arvind and D. Collins, *Phys. Rev. A* **68**, 052301 (2003).
  - [22] L. K. Grover, *Phys. Rev. Lett.* **79**, 325 (1997).
  - [23] C. Zalka, *Phys. Rev. A* **60**, 2746 (1999).
  - [24] J. A. Jones, *Fortschr. Phys.* **48**, 909 (2000).
  - [25] R. Laflamme, D. G. Cory, C. Negrevergne, and L. Viola, *Quantum Information and Quantum Computation* **2**, 166 (2001).
  - [26] M. S. Anwar, D. Blazina, H. A. Carteret, S. B. Duckett, T. K. Halstead, J. A. Jones, C. M. Kozak, and R. J. K. Taylor, *Phys. Rev. Lett.* **93**, 040501 (2004).
  - [27] M. Anwar, D. Blazina, H. Carteret, S. Duckett, and J. Jones, *Chem. Phys. Lett.* **400**, 94 (2004).
  - [28] J. A. Jones, M. Mosca, and R. H. Hansen, *Nature* **393**, 344 (1998).